

Principi, prìncipi e principianti: valutazione ed altro. Franco De Anna

Appartengo certo all'ultima categoria richiamata dal titolo (e mi scuso in anticipo della provocazione). Mi piacerebbe molto, in quanto appartenente a tale categoria, che da qualche parte si sviluppasse un sensato confronto "scientifico" sul tema "rilevazioni nazionali", possibilmente depurato sia dai pregiudizi (politici e non, personali e non, culturali e non, corporativi e non), sia dal primato della "opinione personale" (sulla quale mettere a rendita l'autorevolezza personale), sia infine da qualche "cattedra" più o meno consolidata da meriti cumulati in altri tempi e contesti. Per pura strumentalità faccio riferimento a due "prìncipi" (li considero tali) intervenuti recentemente sul tema: Luciano Canfora e Maurizio Tiriticco. Il primo con dichiarazioni pubbliche, più volte e variamente riprese, contro i test INVALSI perché (riassumo) lontani dall'interpretare lo "spirito critico" il cui sviluppo rappresenta la missione fondamentale dell'istruzione e della formazione. Il secondo con trattazione esplicita e finalizzata sulla "scientificità" delle rilevazioni INVALSI, messe a confronto con una "tassonomia" scientifica di "cosa debbono essere i test". Maurizio Tiriticco elenca una tassonomia consolidata degli strumenti e poi confronta quanto proposto dall'INVALSI con tale tassonomia, per mettere in rilievo gli scostamenti ed i difetti. Nulla da eccepire (almeno in prima istanza). Anzi mi pare un intervento che soddisfa l'istanza posta in apertura: una analisi critico-scientifica sugli strumenti. Anche se non ne condivido le conclusioni. Ma, da principiante, nutro diffidenza istintiva con le tassonomie "scientifiche" e con ciò che rappresentano (lo stato "consolidato" di una disciplina).

Vivo, per mia fortuna, in un luogo circondato da "natura". La strada che porta alla mia casa è contrassegnata da segnali che indicano la presenza di animali selvatici: è facile incontrare cinghiali, fagiani, qualcuno sussurra lupi. Un fagiano contro il parabrezza della mia moto mi fece fare un bel capotombolo in un'alba di qualche tempo fa.

Un conoscente, in auto, ebbe a che fare con un cinghiale in modo cruento (per il cinghiale e per l'auto, per sua fortuna). Si lamentò con me per il fatto che il cartello stradale che invitava alla cautela per la presenza di animali selvatici portasse come contrassegno un cervo. "Che ne sapevo dei cinghiali... non c'era scritto..." si lamentava.

Una evidente non corrispondenza tra la "tassonomia" zoologica e il "test" rappresentato dalla segnaletica stradale, rispetto alla comprensione dell'automobilista. Che valutazione diamo a quest'ultima? Che è sbagliata la segnaletica?

Propongo, ma ricordo che sono un principiante, un altro esempio. Se agli studenti somministrassi una frase come (è un classico e me ne scuso, da principiante) "una donna senza un uomo è nulla", e chiedessi di dare alla frase una punteggiatura adeguata, potrei ricavarne alcuni esiti assolutamente diversi. Per esempio: "Una donna, senza un uomo, è nulla". Ma anche "una donna: senza, un uomo è nulla".

Non c'è, come ovvio, nelle risposte alcun criterio di "correttezza"; ma, accidenti, quanto "rivelano" le risposte....

Naturalmente Maurizio ha assolutamente ragione se l'argomentazione si riferisce all'uso di strumenti di rilevazione degli apprendimenti e rigorosamente utilizzati come tali e con quei limiti. Ma se l'ambizione è quella di avere una rilevazione che non si limiti al puro riscontro dell'apprendimento, ma si cimenti con variabili più complesse come i livelli di competenza, le capacità di problem solving ecc... che significato assume il richiamo a tassonomie consolidate in un tempo il cui il cuore del problema era la mera misura degli apprendimenti?

La mia personale dialettica con Tiriticco sulla questione e suggestione delle "competenze" (che io considero largamente marcata da improprietà, e che invece Maurizio sponsorizza, sia pure con il rigore che lo caratterizza) è reperibile in rete, per il lettore che lo desideri.

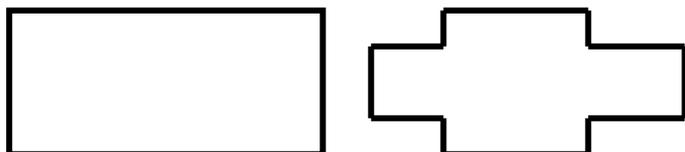
Personalmente ho più volte ricordato che ogni accertamento, non dico valutazione, delle competenze richiede l'esplorazione con strumenti di tipo psicodiagnostico, test compresi. Ma essi sono certo diversi da quelli tassonomizzati da Maurizio e diretti alla misura dell'apprendimento. Altri due esempi tratti (per scrupolo di "distanziamento scientifico") da vecchie rilevazioni. Il primo (PISA di qualche anno fa):

Agli studenti è sottoposta una immagine di una scatola contenente la pasta e su cui era scritto “500 grammi. Cuoce in 10 minuti”. La domanda era “Quanti minuti occorrono per cuocere 250 grammi di pasta?”

Una gran parte (significativa) degli studenti italiani rispose dimezzando i tempi di cottura. Ma la cosa più rilevante fu che andando a controllare le risposte si scoprì che l’errore era stato compiuto soprattutto dai “più bravi”, almeno secondo il giudizio degli insegnanti. Da quelli cioè che avevano “capito” le proporzioni, e avendo “appreso” l’argomento del “programma”, applicavano il loro apprendimento ad ogni occasione, appropriata o meno.

Il secondo (idem).

Sono proposte due figure rappresentanti due “giardini” da recintare.



Le domande: il primo giardino ha lunghezza di 10 metri e larghezza di 5 metri. Quanta rete occorre per recintare?

Quanta rete è necessaria per il secondo?

Alla prima domanda gli studenti italiani rispondono correttamente (sanno calcolare il perimetro di un rettangolo, con “formula” opportunamente “appresa”).

I medesimi studenti alla seconda domanda non sanno rispondere (e anche qui sono i “migliori” nella valutazione dei docenti). E’ evidente che la seconda figura è ottenuta per traslazione di segmenti e che, dunque, il perimetro è il medesimo. Ma i nostri studenti cercano la formula.

E’ improprio sostenere che nell’insegnamento della geometria nelle nostre scuole “forse” c’è un problema? Ma se volessimo sondare la profondità potenziale di un “banale” test, potremmo anche formulare un quesito di grande portata speculativa: “le figure geometriche equivalenti (stessa area), hanno anche uguale perimetro?”.

A Maurizio Tiriticco chiedo in quale delle sue rigorose tassonomie includerebbe i due esempi citati. Sono due esempi di “segnaletica sbagliata”? E con analogo rispetto chiedo all’altro “principe” Luciano Canfora, l’illustre e sempre letto filologo, se per caso non ritenga che anche questi banali esempi non contengano qualche sfida nella applicazione di quello “spirito critico” che giustamente ritiene fondamentale nella “formazione” delle nuove generazioni. O preferiamo scrupolosi alunni che applicano “le formule”?

E alle migliaia di colleghi docenti, certo non principianti come me, chiedo se per caso una lettura critica di “dati” come questi non consenta di sviluppare qualche analisi critica sui modi di insegnare, sulle codificate certezze di alcuni “primati”.

So di avere proposto esempi rudimentali, e forse impropri. Ma credo abbiano almeno la funzione di indicare la necessità da cui sono partito: forse vale la pena di sviluppare una riflessione scientifica e professionale sul senso delle rilevazioni nazionali e, su tale base, delineare proposte, modifiche, miglioramenti. Anche abolizioni, purchè se ne rintraccino i motivi “veri”, scientifici, professionali, e non autodifensivi; o i meccanismi usuali che consistono nel autoconstruirsi i bersagli per essere più sicuri di centrarli nello spunto polemico.

Almeno dalla fine degli anni ’70 (ricerche internazionali IEA) siamo di fronte a responsi sempre ripetuti: i risultati di apprendimento della nostra scuola, nel confronto internazionale, occupano mediamente i gradini più bassi. Ma (e per me la cosa costituisce il problema “vero”) a tale risultato medio corrisponde un livello di grande varianza che testimonia la non equità interna del sistema di istruzione. Una variabilità che corrisponde a stratificazioni geografiche “storiche”, ma anche, all’interno dei medesimi comparti, tra luogo e luogo, tra scuola e scuola, e finanche all’interno di ogni scuola, tra classe e classe...

Ciò era vero in anni (appunto la seconda metà dei '70) nei quali la spesa pubblica per l'istruzione era affluente e costante. Ma non ci curavamo di questi "dati".

Ci siamo per anni accontentati della consolazione che, data la varianza dei risultati, ad una media scadente corrispondesse comunque il fatto che i risultati migliori dei nostri studenti fossero confrontabili con i migliori di quelli degli altri paesi. (La disuguaglianza ci consolava: i nostri migliori erano migliori come gli altri. Guardavamo al primo terzile. Che ci importava del rimanente? Ah!!! Il nostro buon vecchio Liceo Classico... Spero non sia a questo il riferimento di Canfora allo "spirito critico". Continuo a pensare vi sia una sostanziale differenza- e per fortuna! - tra Canfora e la Mastrocola...)

Bene, oggi non è neppure più così... E dove ancora si resiste nel confronto internazionale è dopo i 15 anni di età... nelle fasi terminali della secondaria superiore, dove la popolazione scolastica è ormai "scremata" da ripetuti processi di selezione, affidata non all'INVALSI, ma alla "valutazione dei docenti", loro prerogativa professionale intangibile... Ad essa si deve lo "spiaggiamento" di oltre il 30% della popolazione scolastica, come cetacei abbandonati dalle onde... e quando si tenti la lotta alla dispersione, ad essa si deve il limitarsi ad "abbassare l'asticella".

Certo: no al "teaching to test...". Ma che dire del "teaching to programma"? o "To compito in classe"? o "To interrogazione quadrimestrale"? Certo: da principiante coltivo incertezze su tali verità... No al pericolo conformistico dei test; ma sì a cosa?

I dati mi dicono invece che diamo corso ad un sistema di istruzione scadente e non equo.

E la cultura valutativa espressa dai docenti non è una "risorsa professionale" su cui fare leva, ma un problema da affrontare e risolvere, perché diventi davvero una risorsa.

Se la valutazione di sistema serve a dare contezza e misura di ciò, il problema è assicurarsi che sia la migliore possibile... migliorandola... Se la "valutazione di sistema" fornisce spunti diagnostici anche per una valutazione di livello "micro" (competenze e professionalità dei docenti, organizzazione dei processi) in chiave di possibili interventi migliorativi, meglio ancora...

Assicurarsi che la "filiera valutativa" (cosa, chi, perché e come valutare...) sia la migliore possibile migliorandola diventa obiettivo primario e si spererebbe comune a chi abbia davvero a cuore la scuola.

La ricetta sta nell'aumentare l'investimento in istruzione? Certo... Anche se il riferimento ad anni nei quali la spesa per istruzione era consistente ed affluente pur con risultati scadenti non equi dovrebbe fare riflettere.

Se la macchina che produce l'istruzione è scadente e iniqua possiamo semplicemente pensare di correggerla aumentando le risorse da riversare in essa? Lo fareste mai in una impresa? Significa garantirsi il fallimento..

"Investire" è diverso da "spendere" perché implica uno stringente rapporto con il risultato... Cioè reclama valutazione della redditività dell'investimento. E, trattandosi di risorse pubbliche che provengono dalla fiscalità dei cittadini, tale rendiconto è la condizione essenziale della legittimazione e del consenso sociale, al di là delle "giaculatorie" o degli slogan. E' la democrazia. Per avere risultati occorre investire, certo, ma anche e soprattutto "cambiare la macchina", la sua organizzazione, le sue modalità di funzionamento, il suo modo di impiegare il personale (e, sia chiaro, non sto parlando di "indicazioni" o "programmi", ma di organizzazione...)

Da oltre trent'anni la politica scolastica procede senza alcun strumento di "valutazione della politica pubblica" che non sia esclusivamente limitato al "giudizio politico" (che è cosa diversa).

L'implementazione delle strategie, le condizioni di realizzazione, i risultati, non sono mai stati oggetto di valutazione. Politics e mai policy. Senza malizia: se c'è un avversario "oggettivo" (altra cosa sono le rappresentazioni) della cultura della valutazione estesa al sistema, risiede in Viale Trastevere (e pour cause...). Ma quanti alleati è capace di trovare!!!

Un terzo "principe" interviene sulla valutazione, ed è collettivo: il Sindacato, la CGIL che è per giunta l'organizzazione sindacale cui appartengo (e per biografia assicuro che il termine "appartenere" significa qualche cosa di più di "essere iscritto").

Sulle questioni di fondo tra valutazione e sindacato mi sono espresso estesamente in un recente convegno. Rimando a quel documento (¹)

La CGIL propone un appello da sottoscrivere, che, al di là della articolazione interna delle argomentazioni proposte, ha un impatto immediato “no al sistema nazionale di valutazione”. Certo l’argomentazione è più complessa. Ma, appunto, va ricercata “tra le righe”. Data la richiamata appartenenza, mi costringo a farlo vincendo la tentazione di accartocciare l’appello.

Si rinvia alla necessità di costruire “un vero” sistema di valutazione di sistema, valorizzando autovalutazione ecc... Sensato. Ma sono più di dieci anni che ho partecipato a tutti i tentativi di sperimentazione di modelli di valutazione sistemica, a partire da quelli relativi all’applicazione della “novella” autonomia (il Monitoraggio autonomia dei primi anni 2000) a quelli di valutazione dei Dirigenti Scolastici (le varie versioni SIVADIS). E nel conto metto anche le valutazioni dei progetti PON. Ho esplorato da vicino i tanti tentativi di costruzione di protocolli “autovalutativi” da parte delle scuole. Tanti tentativi e diversi protagonisti. Sperimentazione mai conclusa (valutata). Chiedere in Viale Trastevere i motivi.

Siamo, oggi e attraverso scorci di quadro politico che conosciamo, di fronte ad un decreto che configura un Sistema Nazionale di Valutazione con evidenti difetti e parzialità. Ma è a “maglie larghe”: si presta ad una attenta, realistica e critica operazione di realizzazione e di determinazione dei contenuti con i quali riempire quelle maglie larghe. Questo è il lavoro del Sindacato: contrattazione, formale o meno, negoziazione, contrasto migliorativo, proposta capace di risolvere, mobilitazione di risorse intellettuali e culturali...

Che significa, oggi, rinviare ad un “vero” sistema di valutazione? Che finchè non sia abbia “il meglio” non si parte? E nel frattempo ci rassegniamo alla macchina scadente e non equa? Non ci cimentiamo, anche attraverso la valutazione delle politiche pubbliche, a comprendere come, dove, quanto investire, limitandoci ad invocare investimenti “comunque”?

Mettere a frutto la cultura autovalutativa nel frattempo rielaborata sperimentalmente da almeno un ventennio è certo una indicazione sensata; ma appartiene, appunto alla fase di implementazione, non di definizione di un decreto. E osservazione analoga potrei ripetere per quanto attiene alle diverse sperimentazioni di modelli di valutazione del personale (i Dirigenti). Se non si mette in campo tale lavoro, ai difetti formali del Regolamento si sommano quelli di una implementazione parziale e rudimentale. E mentre il serpente si morde la coda e Viale Trastevere dorme sonni tranquilli, continua il degrado e si abilita, al massimo, l’indignazione frustrata e impotente.

Per gli apprendimenti si vorrebbe una rilevazione campionaria e non censuaria.

Ho predicato e difeso l’approccio campionario in anni non sospetti. Dunque non ho obiezioni di principio. Ma è necessario che la proposta si accompagni a qualche specificazione ulteriore per rintracciarne il senso. La rilevazione campionaria è certamente sufficiente al decisore politico e amministrativo. Ma il “valore politico” (policy) di tale strumento è limitato ai decisori. E’ questo che si vuole?

Può mantenere il valore di “strumento diagnostico” utile ai processi di insegnamento e apprendimento a livello micro (le singole scuole, l’autoanalisi e autodiagnosi del lavoro dei docenti, il miglioramento professionale ed organizzativo delle scuole stesse..), a condizione che le singole scuole sperimentino in parallelo e autonomamente l’uso di strumenti confrontabili con quelli utilizzati nelle rilevazioni campionarie. Siano disponibili e capaci di “autoposizionarsi”.

Anche questa è una prospettiva non certo nuova: era una ipotesi sulla quale si muoveva il vecchio CEDE, che è il padre dell’attuale INVALSI. Non è proibito riproporla, ma occorre chiedersi perché in passato sia sostanzialmente fallita. (ricordo maliziosamente, tra le altre considerazioni, che quel fallimento ha nomi e cognomi).

La condizione di realizzabilità di un modello fondato su rilevazione campionaria a titolarità centrale e parallelo uso autonomo di strumentazione confrontabile per un lavoro coerente di autodiagnosi da parte delle scuole, è costituito dunque, oltre che dalla disponibilità per le scuole di

¹ Rintracciabile on line presso http://www.proteofaresapere.it/urlo/convegnoDS2013_deanna2.pdf

strumenti coerenti, anche dalla promozione e dagli effetti di una forte socializzazione di metodologie, ma soprattutto di significati professionali sul contenuto, il valore e la omogeneità della valutazione utilizzata dai docenti. Dovrebbe cioè coniugarsi con una forte ed autorevole e condivisa operazione culturale sulla “cultura della valutazione” dei docenti. Farla diventare ciò che oggi non è: una risorsa (vedi sopra)

La sfida è dunque, al contrario, se sia possibile e politicamente realistico utilizzare la “forzatura” di un regolamento sul Sistema di Valutazione Nazionale (imperfetto e a maglie larghe, comunque in vigore), e il faticoso maturare di un sistema di rilevazione nazionale dei livelli di apprendimento (che, con grandi contraddizioni, sta comunque consolidando esperienze) come occasione e “campo” delineato sia per una battaglia politica che investe l’implementazione della politica pubblica, la gestione di progetti finalizzati, le strategie di investimento; sia per una battaglia culturale che consolidi la cultura della valutazione e la sua funzione strategica rispetto all’operatore pubblico, a cominciare proprio dalla pubblica istruzione. (Un operatore pubblico sottratto alla sfida della valutazione rappresenta uno degli handicap sostanziali della crisi nazionale e delle disfunzioni e sprechi della Pubblica Amministrazione)

E’ una sfida culturale, scientifica, professionale di grande portata e che ha un orizzonte temporale certo non limitato ad una legislatura o a una maggioranza.

E per finire, la CGIL chiede che non vi siano “prove nazionali” standardizzate negli Esami di Stato. E’ assolutamente corretto chiedere che le prove standard non siano assimilate alle “rilevazioni” sui livelli di apprendimento: le “prove di esame” sono cosa diversa e incommensurabile.

Credo invece che occorrerebbe avere il coraggio di una battaglia per abolire l’Esame di Stato al termine della Scuola Media. Non ha alcun significato istituzionale, con l’obbligo di dieci anni.

Per la secondaria superiore, in costanza di “valore legale” del titolo di studio, quale “ragione scientifica” si oppone all’ipotesi di una prova nazionale? Un tempo al valore certificatorio “nazionale” e “legale” degli Esami di Stato si dava fondamento sulla base della “estraneità” completa delle commissioni (la “terzietà” certificatoria) e sulla estensione complessiva della “rilevazione” valutativa (alla maturità tutte le discipline erano oggetto di valutazione). Non è più così da quarant’anni.

Oggi si dà configurazione “nazionale” alla prova scritta di italiano. Cioè alla prova che più di ogni altra è soggetta a inevitabile “discrezionalismo valutativo”. Si vorrebbe affidare un livello maggiore di “oggettività” alla terza prova (test e su più discipline) ma la si riconduce alla discrezionalità delle commissioni per elaborarla, in evidente contraddizione di finalità e strumenti. Schizofrenia tecnica e politica.

Consiglio alla Direzione della CGIL (il mio sindacato) di consultare le ricerche fatte dall’Accademia della Crusca sulle valutazioni delle prove di italiano e quelle più che istruttive che confrontano la variabilità degli esiti valutativi degli esami di Stato con la variabilità degli esiti dei test di ammissione alle Università. I “controgradienti” relativi alle due valutazioni (i voti più alti della maturità, concentrati in alcune Regioni, sono smentiti clamorosamente dai test di ingresso nelle università della medesime Regioni. . .) sono più che istruttivi per ingigantire i punti interrogativi sull’effettiva equità degli esiti valutativi delle Commissioni degli Esami di Stato. Niente prova nazionale e confidiamo sulle Commissioni? Perché? Per salvaguardarne la “intima” competenza valutativa su cui si fonda la garanzia di equità degli esiti, e dunque il loro “valore legale” e certificatorio?

Guardare i dati da vicino, e senza scongiuri, forse abiliterebbe a qualche diversa rivendicazione. E garantirebbe qualche più tempestiva salvaguardia da misure come quelle prese da qualche “mago della valutazione” in Viale Trastevere (vedi malizia precedente), come il bonus differenziato per i test universitari. . .